

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 07:19:15 ; Search time 16.6794 Seconds
(without alignments)
234.715 Million cell updates/sec

Title: US-09-787-082-9

Perfect score: 119

Sequence: 1 CCSPVCHLEHSLNCTNGG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp-archaea.*

2: sp-bacteria.*

3: sp-fungi.*

4: sp-human.*

5: sp-invertebrate.*

6: sp-mammal.*

7: sp-mhc.*

8: sp-organelle.*

9: sp-phage.*

10: sp-plant.*

11: sp-rodent.*

12: sp-virus.*

13: sp-vertebrate.*

14: sp-unclassified.*

15: sp-rvirus.*

16: sp-bacteriap.*

17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.5	48.3	504	5 Q9XZX8	Q9xzx8 leishmania
2	57.5	48.3	709	5 Q9XZX9	Q9xzx9 leishmania
3	57	47.9	694	5 Q95XU6	Q95xu6 caenorhabdi
4	54	45.4	160	5 Q9XU68	Q9xu68 caenorhabdi
5	51	42.9	418	10 Q948G4	Q948g4 oryza sativ
6	50	42.0	2104	5 Q21281	Q21281 caenorhabdi
7	50	42.0	2104	5 Q964N4	Q964n4 caenorhabdi
8	49.5	41.6	151	5 O17681	O17681 caenorhabdi
9	49	41.2	399	3 O13982	O13982 schizosacch
10	49	41.2	451	13 Q9PT49	Q9pt49 atractaspis
11	49	41.2	555	4 Q96K49	Q96k49 homo sapien
12	49	41.2	569	11 O88716	O88716 mus musculu
13	49	41.2	751	11 Q9R218	Q9r218 mus musculu
14	49	41.2	751	11 Q9QYB2	Q9qyb2 mus musculu
15	48.5	40.8	136	11 Q99PK1	Q99pk1 mus musculu
16	48.5	40.8	643	11 Q9ERV6	Q9erv6 mus musculu

17	48.5	40.8	655	11 Q9WVF5	Q9wvf5 mus musculu
18	48.5	40.8	894	5 Q9U297	Q9u297 caenorhabdi
19	48.5	40.8	1209	11 Q9QX70	Q9qx70 rattus norv
20	48.5	40.8	1210	11 Q9EP98	Q9ep98 mus musculu
21	48	40.3	199	10 Q49095	Q49095 leavenworth
22	48	40.3	199	10 Q49096	Q49096 leavenworth
23	48	40.3	199	10 Q49097	Q49097 leavenworth
24	48	40.3	199	10 Q49098	Q49098 leavenworth
25	48	40.3	199	10 Q49099	Q49099 leavenworth
26	48	40.3	199	10 Q49100	Q49100 leavenworth
27	48	40.3	199	10 Q49101	Q49101 leavenworth
28	48	40.3	199	10 Q49108	Q49108 leavenworth
29	48	40.3	199	10 Q50045	Q50045 leavenworth
30	48	40.3	223	10 Q49111	Q49111 leavenworth
31	48	40.3	226	10 Q9LD74	Q9ld74 arabidopsis
32	48	40.3	238	10 Q9M4B3	Q9m4b3 arabidopsis
33	48	40.3	281	10 Q49113	Q49113 leavenworth
34	48	40.3	322	10 Q49115	Q49115 leavenworth
35	48	40.3	322	10 Q49114	Q49114 leavenworth
36	48	40.3	338	10 Q49109	Q49109 leavenworth
37	48	40.3	347	5 Q8SVW5	Q8svw5 drosophila
38	48	40.3	350	10 Q9ZWK4	Q9zwk4 brassica ol
39	48	40.3	357	10 Q9M4B2	Q9m4b2 arabidopsis
40	48	40.3	359	10 Q9ZWL0	Q9zw10 arabis hirs
41	48	40.3	361	10 Q9LDM3	Q9ldm3 arabidopsis
42	48	40.3	361	10 Q9ZWL4	Q9zw14 arabis stel
43	48	40.3	361	10 Q9ZWL2	Q9zw12 arabis flag
44	48	40.3	361	10 Q9ZWL1	Q9zw11 arabidopsis
45	48	40.3	363	10 Q9ZWK5	Q9zwk5 arabidopsis

ALIGNMENTS

RESULT 1

Q9XZX8 Q9XZX8 PRELIMINARY; PRT; 504 AA.

AC Q9XZX8;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Possible surface antigen.

GN I302.01.

OS Leishmania major.

OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIDLIN;

RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,

RA Rajandream M.A., Barrell B.G.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIDLIN;

RX MEDLINE=98146435; PubMed=9477341;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,

RA Smith D.F.;

RT "A physical map of the Leishmania major Friedlin genome.";

RL Genome Res. 8:135-145(1998).

CC -! COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).

DR EMBL; AL389894; CAC22669.1; -

DR InterPro; IPR000564; 2Fe2S_ferredoxin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR002174; Furin-like.

DR InterPro; IPR002049; Laminin_EGF.

DR SMART; SM00181; EGF; 3.

DR SMART; SM00001; EGF_like; 2.

DR SMART; SM00261; FU; 4.

DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 3.

DR PROSITE; PS01248; LAMININ_TYPE_EGF; UNKNOWN_1.

KW Iron-sulfur.

SQ SEQUENCE 504 AA; 52823 MW; B75450F5FE599263 CRC64;

```

Query Match      48.3%; Score 57.5; DB 5; Length 504;
Best Local Similarity 50.0%; Pred. No. 0.29;
Matches 10; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 2 CSNPV---CHLEHSLNCTNG 18
    ||:| | | | | | | |
Db 398 CSSPTTQPCVEHCNCTCVNG 417

RESULT 2
Q95XU6 PRELIMINARY; PRT; 709 AA.
AC Q95XU6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE Possible surface antigen.
GN L302.02.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL389894; CAC22670.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 709 AA; 73961 MW; C6B1A5348B9B1637 CRC64;

Query Match      48.3%; Score 57.5; DB 5; Length 709;
Best Local Similarity 50.0%; Pred. No. 0.39;
Matches 10; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 2 CSNPV---CHLEHSLNCTNG 18
    ||:| | | | | | | |
Db 603 CSSPTTQPCVEHCNCTCVNG 622

RESULT 3
Q95XU6 PRELIMINARY; PRT; 694 AA.
AC Q95XU6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 76.9 kDa protein.
GN Y61A9LA.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. the C. elegans Sequencing Consortium.";
RL Science 292:2012-2018(1998).

QY 2 CSNPV---CHLEHSLNCTNG 18
    ||:| | | | | | | |
Db 603 CSSPTTQPCVEHCNCTCVNG 622

RESULT 3
Q95XU6 PRELIMINARY; PRT; 694 AA.
AC Q95XU6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 76.9 kDa protein.
GN Y61A9LA.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. the C. elegans Sequencing Consortium.";
RL Science 292:2012-2018(1998).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Ali J., Dempsey S.;
RT "The sequence of C. elegans cosmid Y61A9LA.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024843; AAK70666.2; -.
KW Hypothetical protein.
SQ SEQUENCE 694 AA; 76927 MW; BF1ASC0CC2DB0FDE CRC64;

Query Match      47.9%; Score 57; DB 5; Length 694;
Best Local Similarity 47.1%; Pred. No. 0.46;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CSNPVCHLEHSLNCTNG 18
    | | | | | | | |
Db 634 CRNPICFKHPKECRFG 650

RESULT 4
Q9XU68 PRELIMINARY; PRT; 160 AA.
AC Q9XU68;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T23F6.1 protein.
GN T23F6.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83127; CAB05630.1; -.
SQ SEQUENCE 160 AA; 16334 MW; 9B2D1D0D38E31A7E CRC64;

Query Match      45.4%; Score 54; DB 5; Length 160;
Best Local Similarity 55.6%; Pred. No. 0.38;
Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 2 CSNPVCHLEHSLNCTNG 19
    | | | | | | | |
Db 113 CCQPVC---TNACTNG 126

RESULT 5
Q948G4 PRELIMINARY; PRT; 418 AA.
AC Q948G4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative GATA-1 zinc finger protein.
GN OSJNBA0049012.10.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

OC Ehrhartoidea: Oryzae; Oryza.
 ON NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA Spiegel L., Nascimento L., de la Bastide M., Kirchoff K., Preston R.,
 RA King L., Vil M.D., Baker J., Zutavern T., Santos L., Miller B.,
 RA Kult K., Cunnius D.M., Ballija V., Shah R., Bahret A., Bell M.,
 RA Yang C., Palmer L., O'Shaughnessy A., Dedhia N., McCombie W.R.;
 RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
 OS QJNBA0049012, from chromosome 2, complete sequence."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC069158; AAK98698.1; -;
 DR InterPro; IPR000679; Znf_GATA.
 DR Pfam; PF00320; GATA; 1.
 DR PROSITE; PS0114; GATA_2N_FINGER.2; 1.
 SQ SEQUENCE 418 AA; 4381 MW; 5099191B94A0C609 CRC64;
 Query Match 42.9%; Score 51; DB 10; Length 418;
 Best Local Similarity 57.9%; Pred. No. 2.8;
 Matches 11; Conservative 3; Mismatches 3; Indels 2; Gaps 1;
 QY 3 SNPVCHLEHS--NLCTNGG 19
 DB 177 SSPISVLEHSGFNATNGG 195
 ID Q21281 PRELIMINARY; PRT; 2104 AA.
 AC Q21281;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 230.1 kDa protein.
 GN K07D8.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Favello T.;
 RT "The sequence of C. elegans cosmid K07D8.";
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; LI6679; AAA28092.5; -;
 DR HSP; P02468; ITIE.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR000082; SEA_domain.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF00008; EGF; 20.
 DR Pfam; PF01390; SEA; 4.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00181; EGF; 35.
 DR SMART; SM00179; EGF_Ca; 24.

DR SMART; SM00200; SEA; 4.
 DR SMART; SM00327; vwa; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_15.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_Ca; 3.
 DR PROSITE; PS00024; SEA; 4.
 DR PROSITE; PS00234; VWF_A; 1.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hypothetical protein;
 KW Repeat.
 SQ SEQUENCE 2104 AA; 230051 MW; C98CD462CC434F8B CRC64;
 Query Match 42.0%; Score 50; DB 5; Length 2104;
 Best Local Similarity 41.2%; Pred. No. 17;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 2 CSNPVCHLEHSNLCTNG 18
 DB 1286 CGNGLCHLDLGEVCGV 1302
 ID Q964N4 PRELIMINARY; PRT; 2104 AA.
 AC Q964N4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Transmembrane matrix receptor MUP-4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hong L., Elbl T., Franzini-Armstrong C., Ward J., Rybicka K.K.,
 RA Gatewood B.K., Bucher E.A.;
 RT "MUP-4 is a novel matrix receptor with essential functions in
 RT epithelial cell adhesion at hemidesmosomes and transmission of muscle
 RT forces in Caenorhabditis elegans.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF289202; AAK69172.1; -;
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR000082; SEA_domain.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF00008; EGF; 16.
 DR Pfam; PF01390; SEA; 2.
 DR Pfam; PF00092; vwa; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_13.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
 DR PROSITE; PS01187; EGF_Ca; UNKNOWN_2.
 DR PROSITE; PS00024; SEA; 2.
 DR PROSITE; PS00234; VWF_A; 1.
 KW Receptor.
 SQ SEQUENCE 2104 AA; 230035 MW; 31174C0CCDB61CC2 CRC64;
 Query Match 42.0%; Score 50; DB 5; Length 2104;
 Best Local Similarity 41.2%; Pred. No. 17;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 2 CSNPVCHLEHSNLCTNG 18
 DB 1286 CGNGLCHLDLGEVCGV 1302
 ID Q17681 PRELIMINARY; PRT; 151 AA.
 AC Q17681;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)

```
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE C49F5.5 protein.
GN C49F5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81485; CAB03976.1; -.
DR InterPro; IPR000197; TAG_finger.
DR Pfam; PF02135; zf-TAG; 1.
SQ SEQUENCE 151 AA; 16936 MW; 8B1D37C89586C334 CRC64;

Query Match 41.6%; Score 49.5; DB 5; Length 151;
Best Local Similarity 45.5%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

QY 2 CSNPVC-----HLEHSNLTCTG 18
   | | | | | | | | | |
DB 89 CDPSCGLFKFKYTLHLNMCCTG 110

RESULT 9
ID O13982 PRELIMINARY; PRT; 399 AA.
AC O13982;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical TRP-ASP repeats containing protein C25H1.08C in
DE chromosome I.
GN SPAC25H1.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE BETA TRANSUCIN FAMILY; CONTAINS TRP-
CC ASP DOMAINS.
DR EMBL; Z98944; CAB11604.1; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
DR Hypothetical protein; Repeat; WD repeat.
KW DOMAIN 3
FT DOMAIN 3
SQ SEQUENCE 399 AA; 43431 MW; B536BD0A088D037F CRC64;

Query Match 41.2%; Score 49; DB 3; Length 399;
Best Local Similarity 62.5%; Pred. No. 5.7;
Matches 10; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 4 NPVCHLEHSNLTCTNG 19
   | | | | | | | | | |
DB 70 NPV-----HSNLTASGG 81
```

```
RESULT 10
Q9PT49
ID Q9PT49 PRELIMINARY; PRT; 451 AA.
AC Q9PT49;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metalloproteinase precursor (Fragment).
OS Atractaspis engaddensis (Israeli burrowing asp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Atractaspididae; Atractaspis.
OX NCBI_TaxID=8600;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=VENOM GLAND;
RA Duccancel F.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF186368; AAF01040.1; -.
DR HSSP; P18619; 1FVL.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR00130; Zn_MTpeptdse.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; disintegrin; 1.
DR SMART; SM00030; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Signal.
FT NON_TER 1
FT SIGNAL <1 2 POTENTIAL.
FT CHAIN 31 235 METALLOPROTEINASE.
FT CHAIN 236 328 DISINTEGRIN.
SQ SEQUENCE 451 AA; 50286 MW; 8C1395BF9B9D3EF8 CRC64;

Query Match 41.2%; Score 49; DB 13; Length 451;
Best Local Similarity 38.9%; Pred. No. 6.3;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCNPVCHLEHSNLTCTG 18
   | | | | | | | | | |
DB 269 CCNATTCKLQHCACQDSG 286

RESULT 11
Q96K49
ID Q96K49 PRELIMINARY; PRT; 555 AA.
AC Q96K49;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDNA FLJ14681 fis, clone NT2RP2004270, weakly similar to protein PTM1
DE precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027587; BAB55214.1; -.
SQ SEQUENCE 555 AA; 63536 MW; 7D9B61E5AADF0AA3 CRC64;
```

DR InterPro: IPR000343; Cytochrome bDND;
DR InterPro: IPR002174; Furin-like

```

DR SMART: SM00261; FU: 2.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Receptor.
FT NON_TER 1 1
SQ SEQUENCE 136 AA; 14971 MW; 450A8761345BF1AF CRC64;
Query Match 40.8%; Score 48.5; DB 11; Length 136;
Best Local Similarity 62.5%; Pred. No. 2.6;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
Qy 3 SNPVCHLEHSLCTNG 18
Db 106 ANNVCHLCHAN-CTYG 120

```

Search completed: March 17, 2003, 07:26:25
Job time: 17.6794 secs